

# Using Resolution-Constrained Adaptive Meshes for Traveltime Tomography

Jonathan B. Ajo-Franklin \*

*Earth Resources Laboratory, Department of Earth, Atmospheric, and Planetary  
Sciences, Massachusetts Institute of Technology, E34-566, 42 Carleton St.,  
Cambridge, MA, USA, 02142-1324  
(Formerly Department of Geophysics, Stanford University)*

Jaime A. Urban

*Department of Geophysics, Stanford University, Rm.360, Mitchell Bldg, 397  
Panama Mall, Stanford, CA, USA, 94305*

Jerry M. Harris

*Department of Geophysics, Stanford University, Rm.360, Mitchell Bldg, 397  
Panama Mall, Stanford, CA, USA, 94305*

Reference as :

Ajo-Franklin, J.B., Urban, J.A., and Harris, J.M. 2006.

Using resolution-constrained adaptive meshes for traveltime tomography

Journal of Seismic Exploration, 14: 371-392

---

**Abstract**

We present a travelttime tomography method which exploits a greedy mesh refinement algorithm to construct adaptive parametrizations with guarantees on model resolution. The refinement strategy is implemented on an unstructured trigonal mesh of cells, each with a constant gradient of slowness. We extend previous work on this theme by adding a new class of refinement operations which produce higher quality meshes with fewer poorly formed triangles. Zones of refinement are chosen based on the properties of the diagonal of the model resolution matrix. The resulting mesh guarantees a well-posed inversion problem by construction. Our algorithm is tested on a synthetic crosswell tomography problem with spatially variable angular aperture and a region of negligible ray coverage; the results demonstrate the advantages of adaptive parametrization in comparison to inversion using a regular grid and a single regularization parameter. We also process a large multiwell seismic dataset with irregular ray coverage collected at a shallow site with groundwater contamination.

*Key words:* crosswell seismic methods, travelttime tomography, inverse theory, near-surface geophysics

*PACS:*

---

---

\* Corresponding author.

*Email addresses:* [jfrank@erl.mit.edu](mailto:jfrank@erl.mit.edu) (Jonathan B. Ajo-Franklin), [urban@pangea.stanford.edu](mailto:urban@pangea.stanford.edu) (Jaime A. Urban), [harris@pangea.stanford.edu](mailto:harris@pangea.stanford.edu) (Jerry M. Harris).

## 1 Introduction

Seismic traveltime tomography provides a robust methodology for quantitative reconstruction of P-wave velocities across multiple geological scales ranging from deep mantle imaging (Bijwaard et al., 1998) to reservoir monitoring applications (Harris et al., 1995) and environmental site characterization (Majer et al., 2000) (Daley et al., 2004). The application of traveltime tomography algorithms to crosswell seismic datasets is often complicated by irregular acquisition geometries, variable signal quality, and the underdetermined nature of the resulting inverse problems. Large regions with high levels of seismic attenuation are particularly difficult to effectively image; the loss of signal in the attenuating region often prevents picking all of the acquired traces, resulting in a ray geometries with significant spatial gaps. In addition to large coverage gaps, differences in interwell spacing and S/N levels can produce significant variations in angular aperture between well pairs. If multi-well profiles are treated as a single dataset for the purposes of inversion, the choice of a regular mesh and an appropriate regularization operator becomes difficult. We are forced to either use a fine mesh and deal with a large number of unconstrained parameters or to select a coarse mesh where individual cells are well-constrained but spatial resolution is greatly decreased.

Several approaches exist for dealing with the case of sparse data or variations in resolving power in tomographic imaging. One approach uses a fine discretization of regular pixels to generate an underdetermined inverse problem, similar to the continuous case, and then adds information to this system in the form of a spatially adaptive regularization scheme as implemented by Yi et al. (2003) for the resistivity inversion problem. A second technique extrapolates the data set to full aperture and interpolates all measurements onto a regular geometry thus generating a spatially uniform problem as is done by Li and Nowack (2004). The problem in this case is to predict missing data, a

non-trivial task but one which has seen significant advances through the development of autoregressive signal analysis and optimal prediction-error filters (Fomel and Claerbout, 2003).

We will consider a third approach, the problem of finding an optimal parametrization with spatially varying cell dimensions where mesh quality is judged by the properties of the model resolution matrix. This strategy allows us to construct a mesh adapted around irregular ray-coverage with formal constraints on how different zones of the problem are resolved. Intuitively, regions of space with higher ray densities and more complete angular coverage should allow examination of smaller features, assuming the fundamental wave-theoretic limitations of imaging are honored. We will solve the problem of finding high quality meshes for crosswell seismic experiments assuming straight ray-paths and a consequently linear tomographic problem of the classical form,

$$\mathbf{G} \mathbf{m} = \mathbf{d} \tag{1}$$

where  $\mathbf{G}$  is a ray-theoretic forward modelling operator determined by the problem discretization,  $\mathbf{m}$  is the property model, in this case slowness, and  $\mathbf{d}$  is the observed traveltime vector. Our traveltime tomography method exploits a greedy mesh refinement algorithm to construct adaptive parametrizations with guarantees on the properties of the model resolution matrix,  $\mathbf{R}$ ,

$$\mathbf{R} = \mathbf{G}^{-g} \mathbf{G}. \tag{2}$$

where  $\mathbf{G}^{-g}$  is the natural generalized inverse of the forward operator.  $\mathbf{R}$  can be seen as a filter which shows how the imaging experiment, parametrization, and the choice of  $\mathbf{G}^{-g}$  modify the true model (Menke, 1984), i.e.

$$\mathbf{m}_{est} = \mathbf{R} \mathbf{m}_{true}. \quad (3)$$

Our forward modeling and refinement strategies are implemented on an unstructured trigonal mesh of cells, each with a constant gradient of slowness. We extend previous work on this theme by adding a new class of refinement operations which produce higher quality meshes with fewer poorly formed triangles. The algorithm is tested on both a synthetic reconstruction problem and a large environmental field survey with irregular ray coverage due to localized regions of high seismic attenuation.

## 2 Formulation of the Forward Problem on Unstructured Meshes

Unstructured parameterizations, and trigonal meshes in particular, have several advantages over traditional Cartesian meshes for solution of both the forward and inverse problems. Unstructured meshes can be easily constructed to conform to geological interfaces or surface topography. Accurately representing curved features using a regular Cartesian mesh requires spatial oversampling which increases the cost of evaluating the forward problem. The location of interfaces or mesh density can also be used to introduce prior structural information directly into the parametrization of the inversion. Most importantly, unstructured meshes can be adapted to increase mesh quality where the quality metric might include ray coverage, model resolution, or posteriori covariance. From the perspective of the forward problem, unstructured meshes have an advantage over regular Cartesian representations when large spatial regions are homogeneous. In this situation, a small number of irregular cells can be used to effectively represent an area requiring a larger number of regular pixels. In cases where the density of unstructured control nodes/cells approaches that of a regular Cartesian mesh, the higher book-keeping costs of

irregular grids makes them less attractive for the forward computation.

Several parameterization methods have been considered for use in adaptive tomography including cubic-B splines (Michelini, 1993) (Michelini, 1995), Voronoi polygons (Bohm et al., 2000) (Bohm and Vesnaver, 1999), lumped Cartesian blocks (Bijwaard et al., 1998), and trigonal meshes with constant velocity (Curtis and Snieder, 1997) or squared slowness interpolating functions (Cox, 2004). Cubic-B spline representations have desirable second-order continuity properties but require numerical integration procedures to calculate the Frechet derivatives when solving the inverse problem. Voronoi polygons are attractive for adaptive inversion due to their compact support but greatly complicate ray-tracing due to their variable number of cell edges. Cartesian meshes with cells grouped for the inversion process are computationally efficient for certain types of adaptation but lack the flexibility of truly unstructured parametrizations, particularly for interface representation. Trigonal meshes are a useful building block for tomographic imaging due to their natural relationship to constant gradient interpolating functions. Since a gradient can be uniquely defined by three points, trigonal meshes enable local linear interpolation with continuity at cell boundaries. Although we use straight rays for our formulation, semi-analytic closed forms exist for curved ray paths in trigonal cells with constant gradients in velocity (White, 1989) or the square of slowness (Červený, 1987).

For our parametrization we have chosen to use a trigonal mesh where slowness varies linearly within each cell. In this case, the traveltime path integral and Frechet derivative contributions can be evaluated in closed form. Figure 1 illustrates the key parameters for a triangular cell and the associated control points. The slowness at any point  $\mathbf{r}$  can be written as a summation of  $N$  linear basis functions operating on  $s_j$ , the slowness value corresponding to control node  $j$ ,

$$s(\mathbf{r}) = \sum_{j=1}^N s_j \phi_j(\mathbf{r}), \quad (4)$$

where  $\phi_j$  can be calculated using

$$\phi_j(\mathbf{r}) = \begin{cases} \frac{[(\mathbf{r}-\mathbf{r}_1) \times (\mathbf{r}_2-\mathbf{r}_1)] \cdot \mathbf{e}_3}{[(\mathbf{r}_j-\mathbf{r}_1) \times (\mathbf{r}_2-\mathbf{r}_1)] \cdot \mathbf{e}_3}, & \text{if } \mathbf{r} \text{ is inside of a triangle with vertex } \mathbf{r}_j \\ 0, & \text{otherwise,} \end{cases} \quad (5)$$

where  $\mathbf{r}_j$  is the position vector of node  $j$  and  $\mathbf{r}_1$  and  $\mathbf{r}_2$  are the position vectors of the other two nodes of the triangle containing  $\mathbf{r}$  and  $\mathbf{r}_j$ .  $\mathbf{e}_3$  is the unit vector normal to the plane. Given our slowness interpolation function (equation 4), we can write the tomographic matrix in terms of the  $i$ th ray and the  $j$ th nodal basis function as,

$$\mathbf{G}_{ij} = \int_{\text{ray } i} \phi_j dl, \quad (6)$$

where for our case the raypath is known *a priori* as a straight segment connecting source and receiver. For this class of basis functions, the elements of  $\mathbf{G}_{ij}$  have a simple analytical expression of the form,

$$\mathbf{G}_{ij} = \sum_{\substack{\text{cells that share node } \mathbf{r}_j \\ \text{and are intersected by ray } i}} \|\mathbf{r}_a - \mathbf{r}_b\| \frac{[(\frac{\mathbf{r}_a + \mathbf{r}_b}{2} - \mathbf{r}_1) \times (\mathbf{r}_2 - \mathbf{r}_1)] \cdot \mathbf{e}_3}{[(\mathbf{r}_j - \mathbf{r}_1) \times (\mathbf{r}_2 - \mathbf{r}_1)] \cdot \mathbf{e}_3} \quad (7)$$

where  $\mathbf{r}_a$  and  $\mathbf{r}_b$  are the intersections of the raypath with the cell edges.

## 2.1 Formulation Of The Inverse Problem

If the trigonal mesh geometry is known *a priori*, we may apply the standard techniques of linear inverse theory to obtain a model estimate. In our case we use a least-squares method with an anisotropic first derivative regularization scheme. Using the definition of  $\mathbf{G}_{ij}$  from equation 7 we seek the least-squares solution of

$$\begin{bmatrix} \mathbf{G} \\ \lambda_x \mathbf{D}_x \\ \lambda_z \mathbf{D}_z \end{bmatrix} \mathbf{m} = \begin{bmatrix} \mathbf{d} \\ 0 \\ 0 \end{bmatrix} \quad (8)$$

which has the normal equation

$$\mathbf{m} = (\mathbf{G}^T \mathbf{G} + \lambda_x^2 \mathbf{D}_x^T \mathbf{D}_x + \lambda_z^2 \mathbf{D}_z^T \mathbf{D}_z)^{-1} \mathbf{G}^T \mathbf{d}, \quad (9)$$

where  $\mathbf{m}$  is the slowness model,  $\mathbf{d}$  is the travel time vector and  $\mathbf{D}_x$  and  $\mathbf{D}_z$  are non-uniform directional first derivative operators. The coefficients  $\lambda_x$  and  $\lambda_z$  are regularization parameters in the appropriate direction. A useful secondary variable to consider is the ratio of the  $x$  and  $z$  regularization parameters,  $r_\lambda = \frac{\lambda_x}{\lambda_z}$ , which takes on large values as we bias  $\mathbf{m}$  towards layered models.

$\mathbf{D}_x$  and  $\mathbf{D}_z$  are defined in terms of a mixed notation where subscript  $k$  refers to trigonal cells while  $j$  refers to control nodes. The derivative contribution for the  $j$ th node and the  $k$ th cell is then,

$$\mathbf{D}^{xkj} = \frac{[\mathbf{e}_1 \times (\mathbf{r}_2 - \mathbf{r}_1)] \cdot \mathbf{e}_3}{[(\mathbf{r}_j - \mathbf{r}_1) \times (\mathbf{r}_2 - \mathbf{r}_1)] \cdot \mathbf{e}_3} \quad (10)$$

$$\mathbf{D}^{z_{kj}} = \frac{[\mathbf{e}_2 \times (\mathbf{r}_2 - \mathbf{r}_1)] \cdot \mathbf{e}_3}{[(\mathbf{r}_j - \mathbf{r}_1) \times (\mathbf{r}_2 - \mathbf{r}_1)] \cdot \mathbf{e}_3} \quad (11)$$

where  $\mathbf{e}$  and  $\mathbf{r}$  are the appropriate unit normal and nodal location vectors respectively,

$$\mathbf{e}_1 = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^T \quad (12)$$

$$\mathbf{e}_2 = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}^T \quad (13)$$

$$\mathbf{e}_3 = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^T \quad (14)$$

$$\mathbf{r} = \begin{bmatrix} x & z & 0 \end{bmatrix}^T \quad (15)$$

We use the iterative LSQR algorithm of Paige and Saunders (1982) to solve equation 8. Constant values for  $\lambda_x$  and  $\lambda_z$  are typically chosen by observation since a prior estimate of data variance is often unavailable. Application of the generalized cross validation (GCV) technique (Wahba, 1990) (Aster et al., 2005) for determination of  $\lambda$  values was attempted for several datasets but failed due to the existence of very broad minima on the GCV curve. When applying the GCV method, we assumed a regularization anisotropy ratio,  $r_\lambda$ , and performed the 1-D search in terms of  $\lambda_z$  values with  $\lambda_x$  defined implicitly. Additionally, we know that one component of error within our experimental traveltime data consists of correlated non-Gaussian noise due to picking bias. Since errors of this type break GCVs assumption of Gaussian noise with a fixed variance, this correlated noise is interpreted as a model component and the resulting GCV value for  $\lambda$  is often unrealistically low.

### 3 Metrics For Tomographic Mesh Refinement

The goal of tomographic mesh adaptation is to generate a parametrization with some stipulated property, possibly fulfilling or optimizing a formal quality measure of the resulting inverse problem such as eigenspectrum metrics (Curtis and Snieder, 1997), properties of the model resolution matrix (Cox, 2004), or null-space power (Bohm and Vesnaver, 1999). For larger problems, heuristic measures of mesh quality including cellular hit count (Bijwaard et al., 1998) are often used. In some cases, mesh adaptation is instead driven by structural goals such as the effective representation of geological zones with high velocity gradients (Bohm et al., 2000).

Bohm and Vesnaver (1999) advocate an interpretive approach to mesh adaptation guided by prior knowledge of structure and graphical measures of null space properties. However, the inclusion of manual interpretation steps is both time consuming for large problems and generates a human bias for the geometry in the final mesh. They parametrize the forward and inverse problem in terms of constant velocity Voronoi cells.

Curtis and Snieder (1997) provide a compelling argument for using properties related to problem eigenspectrum and null-space components instead of ray density for determining mesh quality. They consider a  $2 \times 2$  tomography problem with constant slowness cells and the two parameterizations shown in figure 2. If ray density, the number of rays passing through a given cell, is used as a measure of mesh quality then parametrization (A) with hit counts of 2 in both cells is superior to parametrization (B). However, the tomographic system resulting from parametrization (A) has a significant null-space since any perturbation to  $S_1$  and  $S_2$  such that  $\Delta S_1 + \Delta S_2 = 0$  has no effect on either measured traveltimes. Furthermore, either cell  $S_1$  or  $S_2$  can be subdivided into an arbitrary number of horizontal layers without altering ray density. In parametrization (B), each model parameter is uniquely resolved even though

the ray density is lower than in (A).

Curtis and Snieder (1997) advocate the use of a normalized eigenspectrum flatness measure,  $\Omega$ , of the form,

$$\Omega = \frac{1}{\lambda_1} \sum_{i=1}^N \lambda_i, \quad (16)$$

where  $\lambda_i$  is the  $i$ th eigenvalue of operator matrix  $\mathbf{G}$ . This metric can conveniently be computed in  $O(n^2)$  time in comparison to the full SVD of  $\mathbf{G}$  which requires  $O(n^3)$ . However, the  $\Omega$ -metric is intrinsically non-local since all spatial information is encoded in the right eigenvectors for which the eigenvalue spectrum only provides appropriate weighting. Non-local metrics cannot be used for greedy refinement since we have no way of deciding where to add control nodes. Curtis and Snieder (1997) side-step this problem by fixing the number of unknowns *a priori* and executing an expensive global search over all possible model parameterizations. We have found that Curtis' approach, while conceptually attractive, is difficult to apply to significant problems due to the large number of iterations required for the global search to converge.

A related method proposed by Michelini (1993) uses the gradient of data misfit with respect to perturbations in the parametrization to adapt the position of velocity model control points. While Michelini's method converges for simple cases, it is easily trapped by local minimum in parametrization space and additionally is not posed in terms of a mesh quality metric.

#### 4 Mesh Refinement Algorithm

With an approach to tomographic reconstruction on arbitrary trigonal meshes established, we will now consider the problem of automatically generating parameterizations with the appropriate resolution properties. We adopt a greedy

algorithms for tomographic mesh generation, similar to the method presented by Cox (2004), which sequentially add control vertices until the resolution properties of the problem degrade below a threshold value. In contrast to the search methodology of Curtis and Snieder (1997), the mesh generated by greedy algorithms fulfills a quality constraint by construction but does not guarantee a global minimum for a given quality metric. Refinement and coarsening techniques are two possible approaches to greedy mesh adaptation. The refinement approach begins with a coarse mesh and iteratively add points until a threshold criterion is met whereas coarsening techniques begin with a fine mesh and remove low quality nodes until the quality metric is fulfilled. Our technique is a combination of the two approaches. We start with a coarse mesh and sequentially add nodes until no allowable refinements exist. We then prune remaining nodes from the control point set which do not fulfill the resolution criterion. The general algorithm is outlined by the following pseudo-code :

- (1) Start with a coarse mesh enclosing the source/receiver footprint
- (2) While valid refinements exist
  - (a) Compute Delaunay triangulation for the control point set
  - (b) Compute tomography kernel for the triangulation
  - (c) Compute model resolution matrix
  - (d) If even iteration
    - Collect all edges where
      - both endpoints satisfy the resolution criterion ( $\mathbf{R}_{ii} > \mathbf{R}_c$ )
      - the edge satisfy the minimum length requirement ( $L_{12} > L_c$ )
    - Choose  $n$  longest valid edges - add bisection points to the control point set
  - (e) If odd iteration
    - Collect triangles where
      - all nodes satisfy the resolution criterion ( $\mathbf{R}_{ii} > \mathbf{R}_c$ )

- the triangle satisfies the minimum area requirement ( $A_{123} > A_c$ )
  - Choose  $n$  largest valid triangles - add centroids to the control point set.
- (3) While coarsening locations exist
- (a) Compute Delaunay triangulation for the control point set
  - (b) Compute tomography kernel for the triangulation
  - (c) Compute model resolution matrix
  - (d) Scan nodes which do not fulfill the resolution criterion
  - (e) Remove  $m$  worst nodes

Figure 3 illustrates the core refinement loop. At any given step in the refinement process we begin with a set of control points and compute the associated trigonal mesh. The resolution properties of the mesh are evaluated and the location for the next node is chosen using a rule set which we will outline. As mentioned previously, this process is repeated until no valid update locations are available.

One underlying component in the algorithm is the conversion between the set of control points and their associated trigonal mesh. Since a given set of points can have a large number of possible triangulations, we choose the Delaunay triangulation as a unique mapping between the two structures. The Delaunay triangulation, the dual of the Voronoi diagram, is straightforward to compute and is guaranteed to be locally optimum with respect to minimum angle in the mesh (de Berg et al., 1997).

We use a point-by-point estimate of the diagonal elements of  $\mathbf{R}$ , as the central component of our mesh quality metric. The goal in our mesh refinement process is a parametrization where each point is as close as possible to a threshold resolution value,  $\mathbf{R}_c$ , with the hard constraint that  $\mathbf{R}_{ii} \geq \mathbf{R}_c$ . This metric maximizes the spatial density of control points in a given region as long as all

unknowns are resolved to within  $\mathbf{R}_c$ . A possible improvement to our metric might use the off-diagonal components of  $\mathbf{R}$  to provide a threshold in terms of locality.

Computing the mesh quality metric is the most time consuming component of the adaptive refinement algorithm since  $\mathbf{R}$  must be calculated. For problems of small dimensions, the dense SVD can be used for the calculation of  $\mathbf{R}$ . Following Menke (1984), we use the SVD factorization to decompose the operator,  $\mathbf{G}$ , into the matrix product,

$$\mathbf{G} = \mathbf{U} \mathbf{\Lambda} \mathbf{V}^T = \mathbf{U}_p \mathbf{\Lambda}_p \mathbf{V}_p^T \quad (17)$$

where  $\mathbf{U}$  is an  $N \times N$  matrix of eigenvectors spanning the data space,  $\mathbf{\Lambda}$  is the diagonal  $M \times N$  eigenvalue matrix, and  $\mathbf{V}$  is the  $M \times M$  model eigenvector matrix. The subscript  $p$  denotes that the SVD is truncated at the  $p$ th eigenvalue. In this case the truncated SVD model estimate is  $\mathbf{m}_{est} = \mathbf{V}_p \mathbf{\Lambda}_p^{-1} \mathbf{U}_p^T \mathbf{d}$  and  $\mathbf{R}$  can then be written as,

$$\mathbf{R} = \mathbf{G}^{-g} \mathbf{G} = [\mathbf{V}_p \mathbf{\Lambda}_p^{-1} \mathbf{U}_p^T] [\mathbf{U}_p \mathbf{\Lambda}_p \mathbf{V}_p^T] = \mathbf{V}_p \mathbf{V}_p^T. \quad (18)$$

The choice of the truncation point,  $p$ , is taken as a prior parameter; we tend to choose a large  $p$  value to approximate the unregularized problem. Figure 4 provides a visual example of  $diag(\mathbf{R})$  for a set of crosswell ray paths and a refined mesh with a quality constraint of  $\mathbf{R}_c = 0.1$ .

The primary difference between our approach and the method outlined by Cox (2004) is the fashion in which points are added during the refinement process. At each iteration, the Cox algorithm adds control nodes at the centroid of all triangles where the  $\mathbf{R}_{ii}$  values of the bounding nodes are greater than  $\mathbf{R}_c$ . Part A of figure 5 shows a single step of centroid refinement and the result of

re-meshing with the additional control points. Within this refinement scheme no points will ever be added to the boundary of the domain resulting in long skinny triangles at the edge of the parametrization if too coarse a starting grid is chosen. This effect is particularly problematic in crosswell scenarios where a paucity of points at the domain boundary make the addition of prior information from well logs difficult.

A second possible approach is to add points so as to bisect triangle edges where both bounding nodes fulfill the  $\mathbf{R}_c$  criterion. This approach allows control nodes to be added on the domain boundary resulting in a parametrization which is not biased towards refinement on the interior. Part B of figure 5 depicts an edge bisection refinement step. A failing of the edge bisection update is the generation of preferential alignment of triangles in the end model.

We alternate between the two updating schemes at each iteration to produce meshes without preferential point distribution or preferential triangle alignment. Figure 6 shows the result of centroid refinement (A), edge bisection refinement (B), and the alternating scheme (C).

In addition to the  $\mathbf{R}_c$  criterion, we also apply a size constraint to the point update process. Triangles below a minimum area and edges below a minimum length are not refined. The motivation for the size constraint is to avoid parameterizations which include many triangles far below the resolving power of finite bandwidth field data. In this case, resolving power refers to the smallest feature which can be successfully imaged using the full waveform and not to the elements of  $\mathbf{R}$  which are based on our ray-theoretic forward operator  $\mathbf{G}$ . Since our inversion scheme is based on ray-theory, an infinite bandwidth approximation to the wave equation, no notion of wave-theoretic resolution is present in the actual tomography scheme. In reality, spatial resolution in travelttime transmission tomography is limited by both wavelength ( $\lambda$ ) and the experimental geometry and is variable within a given imaging plane. Al-

though there is no closed form for the spatially varying wave-theoretic resolution, Williamson and Worthington (1993) present a rough estimate of the maximum resolving power ( $r_{min}$ ) of transmission data using Rytov theory,

$$r_{min} \sim \sqrt{L \lambda} \quad (19)$$

where  $L$  is the length of the transmission path. We typically choose minimum edge lengths on the order of  $r_{min}/4$ . This limit is not intended to be a true estimate of the wave-theoretic resolving power of the dataset, but merely a rough lower bound to prevent refinement far below a reasonable length scale.

Another relevant aspect of the updating scheme is choosing the number of control points ( $m$ ) to add or remove at each iteration of refinement. For large values of  $m$ , the refinement algorithm converges quickly with few required calculations of  $diag(\mathbf{R})$ . However, we have found that adding a moderate number of points at each iteration generates meshes with a slightly higher quality, possibly because the sequential updating of  $\mathbf{R}$  allows better choice of point locations.

The coarsening stage is handled in a similar fashion to the mesh refinement. If a given node has a  $\mathbf{R}_{ii}$  value less than  $\mathbf{R}_c$ , that node is a candidate to be removed during the coarsening phase. At each iteration of coarsening we remove  $n$  poorly resolved nodes and then recompute the resolution metric. Since only  $n$  nodes are removed at each step, some freedom exists as to which part of the model nodes are initially removed from. We use a heuristic distance measure to preferentially cull points from the domain interior in early iterations, however we have found that final mesh is generally insensitive to the choice of removal order. After the mesh refinement and coarsening steps are complete, the inversion algorithm described in section 2 is applied to generate the final tomogram.

## 5 A Synthetic Application

Our refinement and inversion algorithm was first tested on a synthetic cross-well dataset with an irregular source/receiver configuration. Panels (a) and (b) of Figure 7 depict the true velocity model used for the synthetic test and the ray path set respectively. The sparse ray geometry has a spatially variable angular aperture, a zone with no transmitted arrivals between 6 and 7 m depth, and several missing source and receiver locations. Synthetic traveltimes were generated using a finely discretized regular mesh and the same straight ray algorithm described in the previous sections. Gaussian noise with a 2% magnitude was added to the synthetic traveltimes. Some features in the true velocity model will not be visible in any tomogram given the sparse source/receiver geometry. In these regions, particularly the coverage gap between 6 and 7 meters depth, the most we can hope for is a clean interpolation between bounding slowness values.

Three tomograms were calculated for this dataset, one adaptive mesh inversion and two regular mesh inversions with different  $\Delta x$  and  $\Delta z$  spacing values (see figure 7). Panel (c) depicts the adaptive mesh generated by the algorithm described previously. For the adaptive meshing phase, the resolution threshold was set to a low value of 0.1 and a minimum edge length of 0.3 m was used. The regular mesh inversions were performed on both fine and coarse parameterizations. Panel (e) depicts the tomogram for a  $20 \times 75$  sample regular mesh while Panel (f) shows the equivalent results for a  $10 \times 40$  regular mesh. The number of unknowns in the inversions were 744, 1500, and 400 for the adaptive, fine regular, and coarse regular meshes respectively. For both regular mesh inversions, control points with no constraining rays were discarded. When these control points were included in the inversion, the results were not usable due to overwhelming image artifacts. As can be seen in panel (c) of figure 7, zones with the densest ray coverage and the highest model resolu-

tion values are represented with a fine parametrization while zones with poor coverage are appropriately coarsened. As can be seen in the tomograms, the adaptive scheme both gracefully handles the variations in ray coverage and does a better job of recovering velocity boundaries. Improvements are visible in the small high velocity lens at 3.5 m depth and the high velocity feature at the bottom of the model.

Figure 8 shows ray-hit and  $diag(\mathbf{R})$  maps for the same three meshes. As can be seen from panels (a) and (d), although the adaptive mesh fulfills our resolution constraints, neither ray hits nor  $diag(\mathbf{R})$  values are uniform. Since the refinement and sparsening operations are intrinsically discrete, the greedy scheme does not have sufficient flexibility to generate parameterizations where the quality metrics of triangles are smoothly distributed. One possible remedy to this problem is the use of a second round of mesh optimization which operates by perturbing nodal locations rather than adding or subtracting nodes.

## 6 Application To A Multi-offset Crosswell Profile

The dataset which drove the development of our adaptive tomography approach was a large multi-well crosswell seismic survey (Geller et al., 2002) (Ajo-Franklin et al., 2002) acquired at the Pinellas DOE facility, a site with confirmed dense non-aqueous phase liquid (DNAPL) contamination. The seismic dataset was acquired as a sequential curtain of seven crosswell surveys extending from a region of the site without free-phase contaminants to a zone where DNAPLs were observed from borehole water sampling. By using a sensitive 24 level hydrophone string and a small piezoelectric source, data was acquired with a center frequency in excess of 5 kHz. One of our processing goals was to jointly invert the entire dataset to allow more effective interpretation of lateral velocity variations. Lateral well spacings varied between 2 and 6 m with vertical source and receiver spacings of either 0.125 or 0.25 m.

The most challenging aspect of the Pinellas seismic dataset was a zone of high seismic attenuation which prevented us from picking first arrivals in a large region of the survey. Panel (A) of figure 9 shows the map of ray-paths with sufficient S/N levels to pick. The resulting source/receiver geometry was highly irregular with differences in both vertical coverage and angular aperture between consecutive well pairs. Panels (B) and (C) of figure 9 show the adaptive mesh generated by our algorithm and the resulting  $V_p$  tomogram. In particular, the M34D-G18 profile had very limited angular aperture, large regions without ray coverage, and low S/N in comparison to the M31D-G20 and G19-G20 sections. For the adaptive meshing step, the resolution threshold was set to a relatively low value of 0.1 and a minimum edge length of 0.2 m was used to prevent meshing below  $\sqrt{L} \lambda/4$ . In total, 4632 picks were used for an inversion with 1419 control nodes.

Visible in panel (B) of Figure 9, our mesh generation algorithm avoided over-parametrization of the low coverage region and provided a velocity estimate in these zones consistent with data recorded above and below. In regions with sufficient coverage, the mesh is refined down to our prespecified limits. Like the synthetic examples, the values for the  $diag(\mathbf{R})$  metric are not uniform across the tomogram with the worst values present in the lower regions with poor angular aperture including the lower portions of the M34D-G18 and G17-G16 sections. Since triangles are not aligned at profile boundaries, some information from the adjacent wells is used in poor coverage regions; in the case of G17-G16, the dense coverage on the boundary with G17 constrains the region between -2 and -3 m elevation. The observed velocity variations for this dataset were on the order of 200 m/s or approximately 11 % of the peak velocity; the tomogram would probably benefit somewhat from the inclusion of ray curvature.

## 7 Conclusion And Future Work

Our current adaptive tomography strategy is limited to relatively small problems with model and data dimensions of less than  $10^4$ , more than sufficient for the datasets acquired at small environmental sites but trivial in comparison to large crosswell or VSP surveys with dimensions on the order of  $10^6$  or 3D reflection surveys with dimensions upward of  $10^8$ . This limitation is largely due to our computation of the SVD of the operator matrix  $\mathbf{G}$  at each of several hundred refinement steps. Each SVD has an asymptotic cost of  $O(n^3)$  in both memory and computational time which makes scaling the algorithm to larger problems difficult. This problem could be partially eliminated through use of an iterative scheme for computing elements of  $\mathbf{R}$ . Cox (2004) uses an adaptation of the LSQR algorithm (Paige and Saunders, 1982) to calculate components of  $\mathbf{R}$  without having to explicitly form  $\mathbf{G}$  while Fomel et al. (2002) uses a similar modification to the traditional conjugate gradient method. In both cases only the action of  $\mathbf{G}$  and  $\mathbf{G}^T$  on  $\mathbf{m}$  are required for implementation. Even with an iterative scheme in place, the cost of computing even a few elements of  $\mathbf{R}$  is still expensive for large problems. Another possible alternative is to develop a cheap mesh quality heuristic based on a combination of ray density and local angular coverage. In this case, angular coverage might provide a secondary constraint to help prevent excessive refinement in cases like geometry (A) of figure 2.

A second limitation of the current scheme is our inability to include ray curvature and therefore handle the full non-linear problem. While we have written several codes capable of calculating curved raypaths on regular Cartesian grids, the extension of these techniques to unstructured trigonal meshes is somewhat more complex. We are in the process of developing raytracing schemes based on the semi-analytic closed form for rays within trigonal cells with a constant gradient of slowness squared (Červený, 1987) and the wavefront construction

algorithm (Vinje et al., 1993). Even after completion of a curved ray tracer for trigonal or tetrahedral meshes, we still must contend with the problem of handling mesh adaptation at each gradient step in the non-linear problem; modification of ray geometry during convergence will force some iterative re-adaptation of the mesh. One approach to solving this problem would be to only add a small number of additional control nodes at each step in the non-linear problem, a strategy analogous to methods which successively relax model-space constraints such as the dynamic smoothing technique proposed by Nemeth et al. (1997).

A more fundamental limitation of the described approach is our reliance on ray theory, an infinite bandwidth approximation, within the reconstruction process. As mentioned previously, the resolution criterion,  $\mathbf{R}_c$ , used in our refinement algorithm is only a measure of how well parameters are resolved within a ray-theoretic context. If the same type of technique were developed within the framework of wavefield tomography,  $\mathbf{R}$  would include the effects of finite bandwidth field data and correspond to a physical resolution rather than simply being a measure of the inverse problem quality.

In summary, we have succeeded in developing an adaptive traveltime tomography algorithm based on greedy mesh refinement. Our formulation provides formal guarantees on model resolution and can be easily applied to problems with irregular source/receiver geometries. We have improved upon previous adaptation schemes by incorporating a new mesh refinement operator which reduces the number of high aspect ratio triangles. Using a complex synthetic test case, we have demonstrated that our algorithm both improves tomogram quality in regions of good coverage and provides consistent property estimates in regions with insufficient data. We have also successfully inverted a substantial field dataset, the Pinellas M31D-G16 multi-offset crosswell seismic profile.

## 8 Acknowledgments

The field experiments referenced in this paper were initiated and sustained by Jil T. Geller of Lawrence Berkeley National Laboratory. Components of this research were supported by US DOE grant # DE-AC-03-76F0098 (PIs Jil Geller and Ernie Majer) under the Subsurface Contamination Focus Area of the Environmental Management Program. The first author was funded through the US EPA's STAR Fellowship program and the Chair's Fellowship for Intergroup Research (Stanford University). We would also like to thank Prof. Nafi Toksöz, and the Earth Resources Laboratory for additional support.

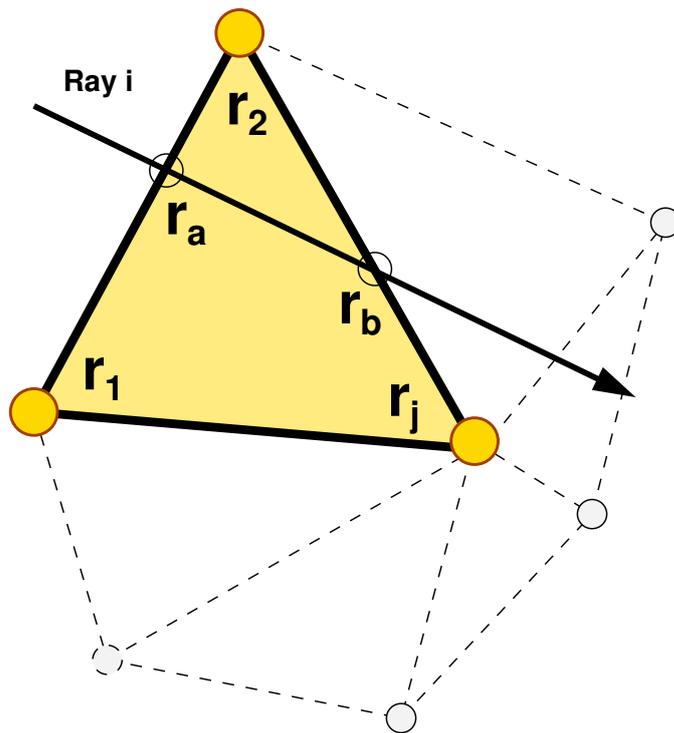


Fig. 1. Illustration of the parameters used for the trigonal model relevant to equations 4, 5, 6, and 7.

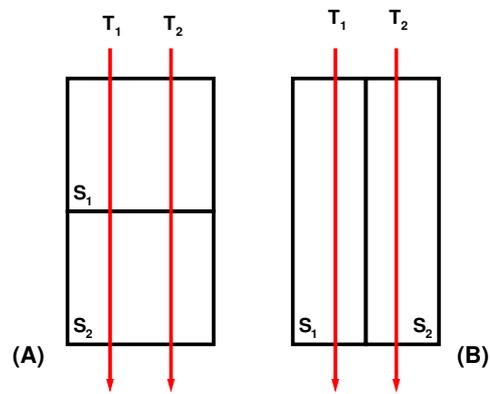


Fig. 2. Ray density vs. Resolution : This trivial example from Curtis and Snieder (1997) demonstrates that a model with a higher cell-by-cell hit-count (A) can have a larger null-space than a comparable model with a lower hit-count (B).

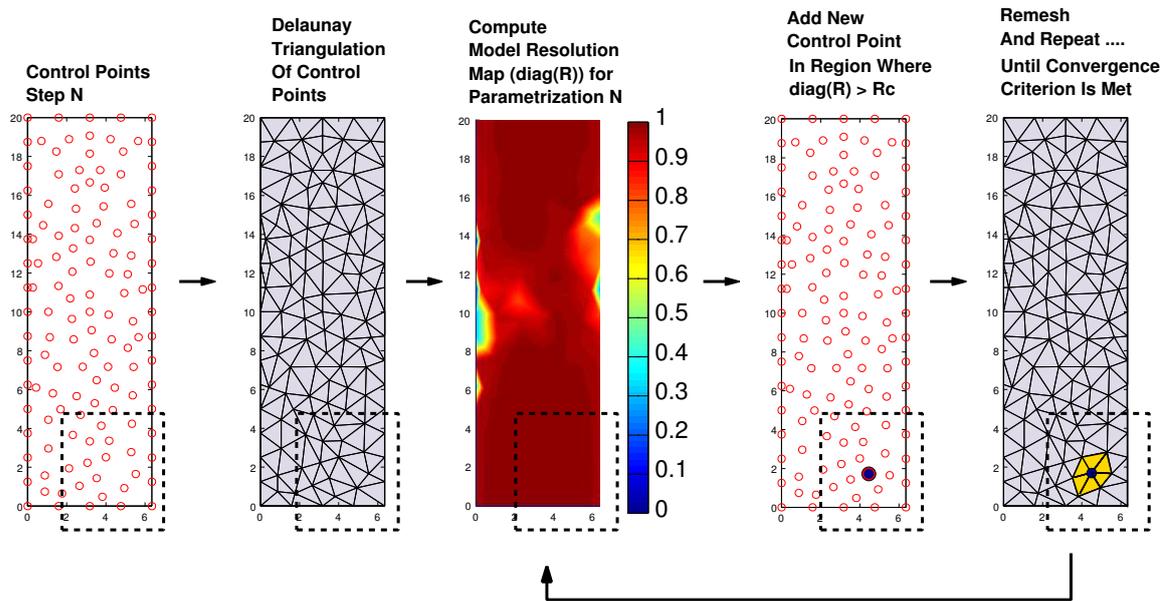


Fig. 3. A schematic view of the mesh refinement process : Starting with a set of control points at iteration  $n$  we calculate a Delaunay triangulation, compute the associated model resolution metric, add another control point and repeat this process until no valid updates exist.

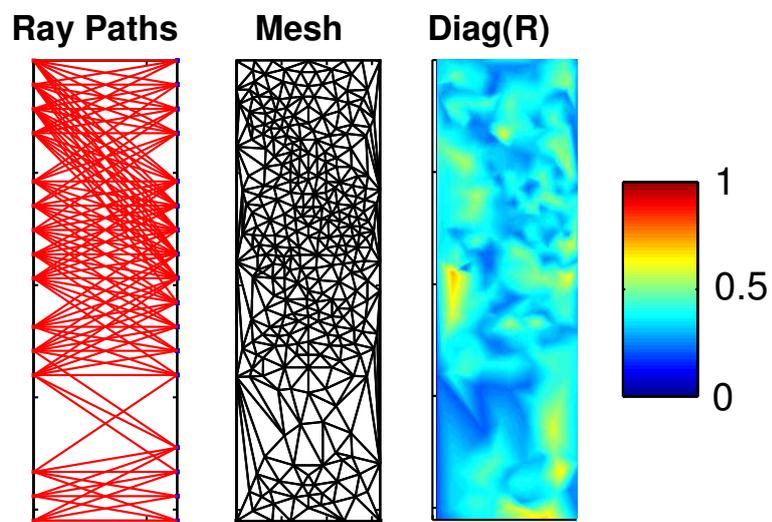


Fig. 4. An example of mesh refinement and the corresponding values for  $diag(\mathbf{R})$  : The left most panel shows a crosswell S/R/ geometry and a set of raypaths. The center panel shows one possible refined mesh with the right panel providing a map of  $diag(\mathbf{R})$  for this particular parametrization.

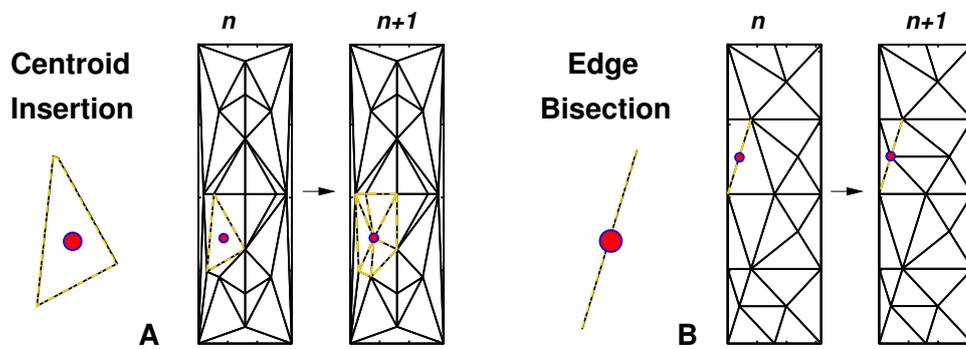


Fig. 5. Two control point updating schemes. (A) centroid refinement, of Cox (2004) (B) edge bisection refinement.

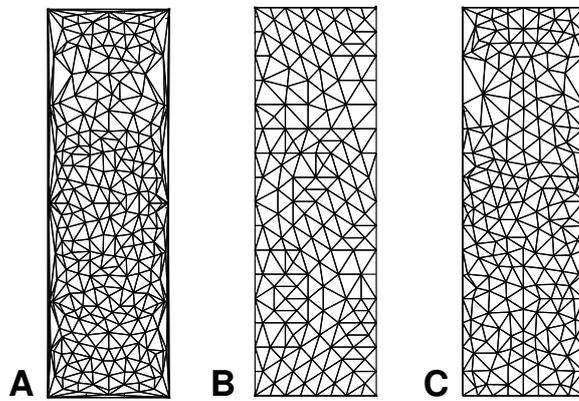


Fig. 6. Meshes generated by the three insertion algorithms. (A) centroid refinement of Cox (2004), (B) edge bisection refinement, (C) alternating update method.

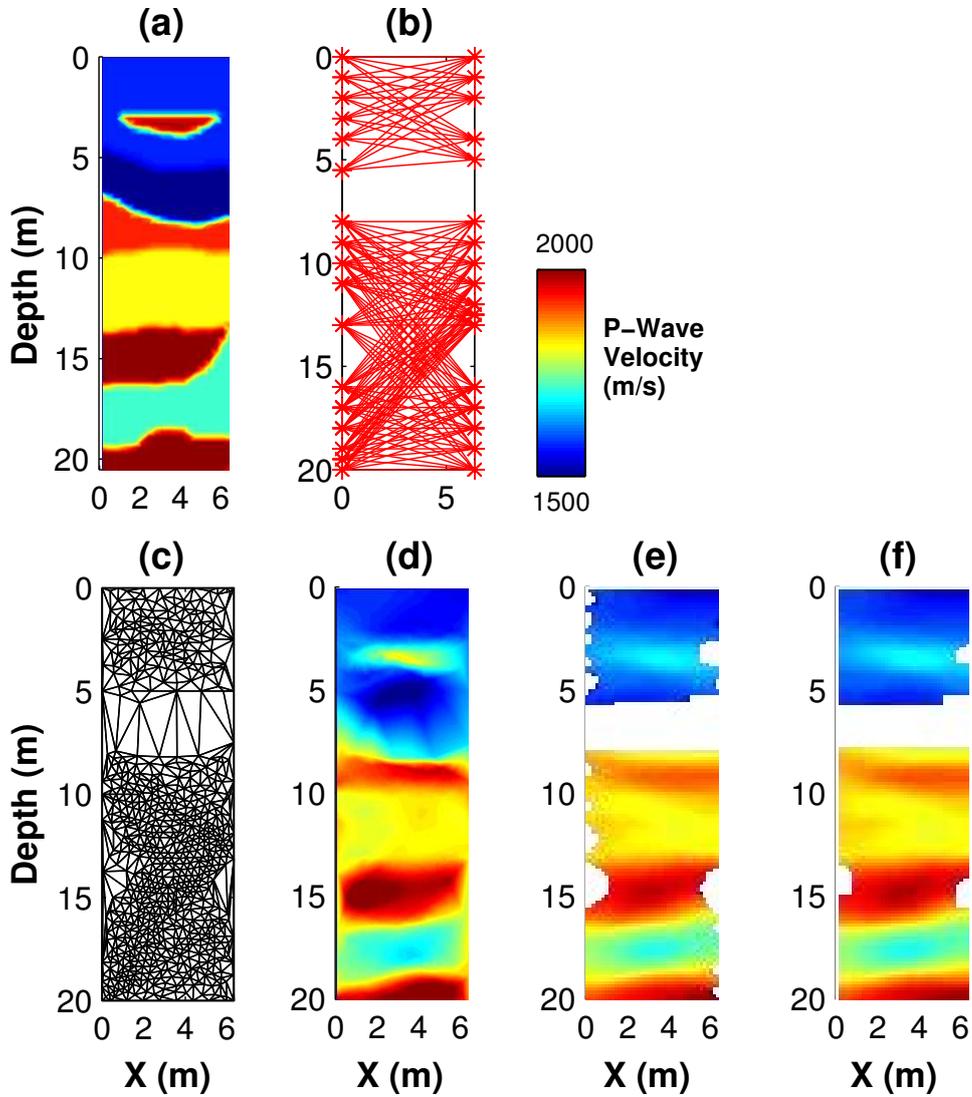


Fig. 7. A synthetic example of our adaptive tomography scheme : (a) the true velocity model, (b) the ray paths used for the inversion, (c) the optimal mesh obtained from our greedy updating scheme, (d) the tomogram calculated using mesh c, (e) the tomogram calculated using a fine ( $20 \times 75$ ) regular mesh, (f) the tomogram calculated using a coarse ( $10 \times 40$ ) regular mesh. The white regions in (e) and (f) correspond to parameters which were not inverted for.

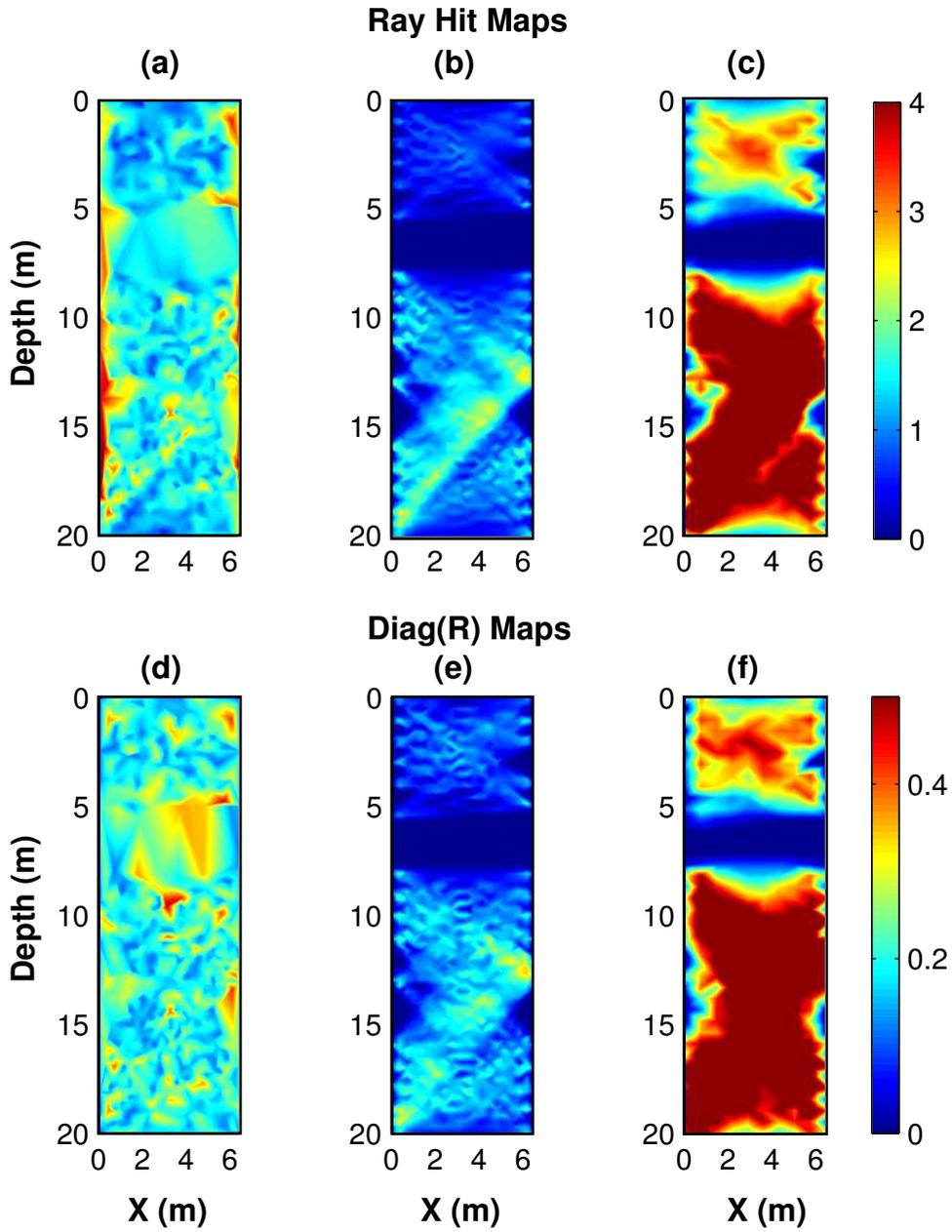


Fig. 8. Maps of ray coverage and resolution properties for the synthetic example : the top row shows the ray hit maps for the adaptive (a), fine regular mesh (b), and coarse regular mesh (c) with a clip of 4 hits applied. The bottom row shows  $diag(\mathbf{R})$  for the same three cases with a clip of 0.5.

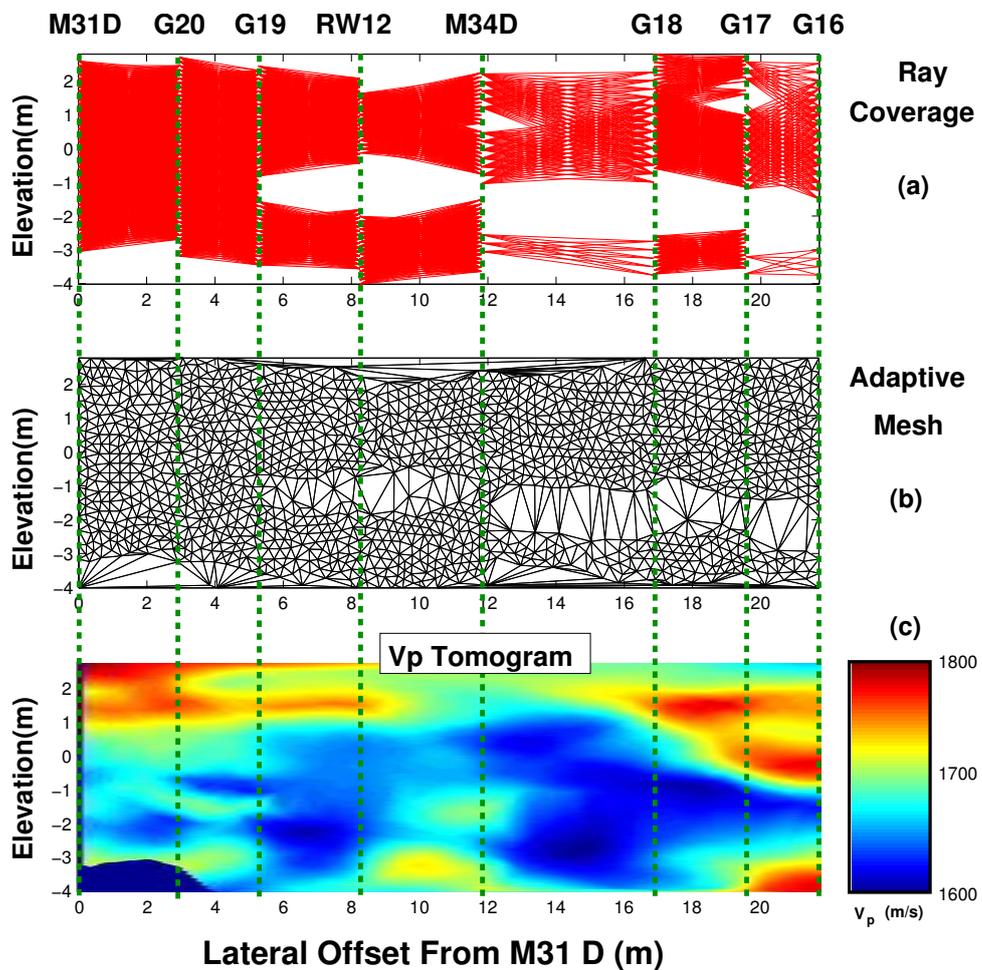


Fig. 9. Multiwell seismic tomography results for the M31D-G16 profile at the Pinellas DOE site : Panel (A) shows a map of picked arrivals in the profile while (B) and (C) depict the adaptive trigonal mesh and the final  $V_p$  tomogram respectively.

## References

- Ajo-Franklin, J., Geller, J., Majer, E., Peterson, J., Williams, K., Harris, J., 2002. Integrated geophysical characterization of a NAPL-contaminated site using borehole and laboratory measurements : T22b-1142. In: EOS Trans. AGU Fall Meet. Suppl. Vol. 83. American Geophysical Union.
- Aster, R., Borchers, B., Thurber, C., 2005. Parameter Estimation and Inverse Problems.
- de Berg, M., van Kreveld, M., Overmars, M., Schwarzkopf, O., 1997. Computational Geometry : Algorithms and Applications. Springer Verlag.
- Bijwaard, H., Spakman, W., Engdahl, E., December 1998. Closing the gap between regional and global travelttime tomography. *Journal of Geophysical Research* 103: 30055–30078.
- Bohm, G., Galuppo, P., Vesnaver, A., 2000. 3D adaptive tomography using Delaunay triangles and Voronoi polygons. *Geophysical Prospecting* 48: 723–744.
- Bohm, G., Vesnaver, A., 1999. In quest of the grid. *Geophysics* 64 (4): 116–1125.
- Cox, B. E., 2004. Tomographic inversion of focusing operators. Ph.D. thesis, Delft University of Technology.
- Curtis, A., Snieder, R., 1997. Reconditioning inverse problems using the genetic algorithm and revised parametrization. *Geophysics* 62 (4): 524–1532.
- Daley, T., Majer, E., Peterson, J., January-February 2004. Crosswell seismic imaging in a contaminated basalt aquifer. *Geophysics* 69 (1): 16–24.
- Fomel, S., Berryman, J., Clapp, R., Prucha, M., 2002. Iterative resolution estimation in least-squares Kirchoff migration. *Geophysical Prospecting* 50: 577–588.
- Fomel, S., Claerbout, J., 2003. Multidimensional recursive filter preconditioning in geophysical estimation problems. *Geophysics* 68: 577.
- Geller, J., Peterson, J., Williams, K., Ajo-Franklin, J., Majer, E., 2002. First

- field test of NAPL detection with high resolution borehole seismic imaging. In: 9th Biennial International Conference on Nuclear and Hazardous Waste Management.
- Harris, J., Nolen-Hoeksema, R., Langan, R., Schaak, . M. V., Lazaratos, S., III, J. R., 1995. High-resolution crosswell imaging of a West Texas carbonate reservoir: Part 1- Project summary and interpretation. *Geophysics* 60: 667–681.
- Li, C., Nowack, R., August 2004. Application of autoregressive extrapolation to seismic tomography. *Bulletin of the Seismological Society Of America* 94 (4): 1456–1466.
- Majer, E., Peterson, J., Williams, K., Daley, T., Gee, G., September 2000. High resolution imaging of vadose zone transport using crosswell radar and seismic methods. Tech. Rep. PNNL-13791, Pacific Northwest National Laboratory, Richland, Washington 99352.
- Menke, W., 1984. *Geophysical Data Analysis : Discrete Inverse Theory*. Academic Press.
- Micheline, A., 1993. Velocity model inversion using parametric curves. *Geophysical Journal International* 115: 337–343.
- Micheline, A., 1995. An adaptive-grid formalism for traveltimes tomography. *Geophys. J. Int.* 121: 489–510.
- Nemeth, T., Normark, E., Qin, F., 1997. Dynamic smoothing in crosswell traveltimes tomography. *Geophysics* 62 (1): 168–176.
- Paige, C., Saunders, M., 1982. An algorithm for sparse linear equations and sparse least squares. *ACM Transactions in Mathematical Software* 8: 43–71.
- Červený, V., 1987. Ray tracing algorithms in three-dimensional laterally varying layered structures. In: Nolet, G. (Ed.), *Seismic Tomography : With Applications In Global Seismology and Exploration Geophysics*. D. Reidel Publishing Company, pp. 99–133.
- Vinje, V., Iverson, E., Gjoystdal, H., 1993. Traveltimes and amplitude estimation using wavefront construction. *Geophysics* 58: 1157–1166.

- Wahba, G., 1990. Spline Models for Observational Data. Vol. 59 of CBMS-NSF Regional Conference Series in Applied Mathematics. Society For Industrial And Applied Mathematics.
- White, D., 1989. Two-dimensional seismic refraction tomography. *Geophysical Journal International* 97: 223–245.
- Williamson, P., Worthington, M., May 1993. Resolution limits in ray tomography due to wave behavior: Numerical experiments. *Geophysics* 58 (5): 727–735.
- Yi, M., Kim, J., Chung, S., 2003. Enhancing the resolving power of least-squares with active constraint balancing. *Geophysics* 68: 931.